UNCLASSIFIED

AD NUMBER
AD828011
LIMITATION CHANGES
TO: **Approved for public release; distribution is unlimited.**
FROM: **Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; DEC 1967. Other requests shall be referred to Office of Naval Research, Arlington, VA 22203.**
AUTHORITY
ONR ltr 27 Jul 1971

THIS PAGE IS UNCLASSIFIED

STANFORD UNIVERSITY

CENTER FOR SYSTEMS RESEARCH

Learning with a Lack of Prior Data

by

Bruno O. Shubert

Technical Report No. 6151-4

Systems Theory Laboratory

BLANK PAGES
IN THIS
DOCUMENT
WERE NOT
FILMED

SEL-67-107

LEARNING WITH A LACK OF PRIOR DATA

by

Bruno O. Shubert

December 1967

Technical Report No. 6151-4

This work was supported wholly by the Joint Services Electronics Programs U.S. Army, U.S. Navy, and U.S. Air Force Under Contract Nonr-225(83), NR 373 360

Systems Techniques Laboratory
Stanford Electronics Laboratories
Stanford University Stanford, California

CONTENTS

I. INTRODUCTION

A convenient model for learning is provided by the sequential compound decision problem of mathematical statistics (see, e.g., [5]). The decision-maker observes a sequence of independent random variables (samples) X_n, the distribution of which varies arbitrarily along the sequence. The decision-maker is required to make a sequence of decisions d_n from a given set of decisions D incurring a loss $L(\theta_n, d_n)$ if his decision at the n-th step was d_n and the sample X_n was distributed according to ν_{θ_n}, where $\theta_n \in \Theta$ is a parameter of the family of distributions $\{\nu_\theta : \theta \in \Theta\}$. The observation X_n thus conveys some information about the parameter θ_n. However, since the decision-maker does not know the distribution of X_n beforehand, he tries to learn during the sequence how to utilize this information to minimize his losses.

A natural criterion of his performance is the average loss during the sequence or the expectation of the latter (average risk). It has been argued (see, e.g., [8]) that, in a sense, the best he can expect is to reduce this quantity to the Bayes risk ρ of the underlying generic decision problem evaluated for the hypothetical prior distribution on Θ equal to the empirical distribution of the sequence of parameters $\{\theta_n\}$. Therefore, the decision-maker should try to reduce the excess of the average risk over the corresponding Bayes risk--the so-called regret R_n--to zero as the index of the sequence $n \to +\infty$. Moreover, since the parameter $\theta \in \Theta$ may vary arbitrarily with n, he should do this uniformly in all sequences $\{\theta_n\}$.

Several papers have been concerned with the question of finding a decision rule for the decision-maker to achieve this goal. If we leave aside those in which some restrictions were imposed on the possible sequences of parameters (usually assuming that $\{\theta_n\}$ is a sequence of i.i.d. random variables), we can divide these works according to the assumptions made about the information available to the decision-maker. This information may be

(1) The initial information concerning the data of the generic decision problem, viz., the family of distributions $\{\nu_\theta : \theta \in \Theta\}$, the loss function L, and properties of the sample space Ξ and the sets Θ, D.

(2) The learned information received gradually during the sequence of decisions. Here, we can distinguish three main cases:

(a) after each decision d_n is made, the decision-maker is told the value of the parameter θ_n;

(b) after each decision d_n, the decision-maker observes a random estimate of the value θ_n;

(c) the decision-maker is not told anything and has to rely only on observations of the samples X_n.

Thus, for example, the case (2a) has been studied by J. Hannan [3], the case (2b) by S. Jílovec and the author [2], the case (2c) by E. Samuel [7] and J. Van Ryzin [5], all assuming complete initial information (all the data known). Of those assuming only partial initial information, namely the family $\{\nu_\theta : \theta \in \Theta\}$ not known, let us mention J. Van Ryzin [6] for the case (2a) and N. Alēns and T. M. Cover [1] for the case (2c), the latter, however, for a slightly modified problem (compound decision problem only).

In this paper, we will make a completely different pair of assumptions, which, in our opinion, may be more appropriate for some learning models. We will assume:

(1) The decision-maker knows only his own decision space D and the sample space Ξ. He knows nothing about the family $\{\nu_\theta : \theta \in \Theta\}$ or the set Θ (which may be infinite), nor does he know the loss function L. The loss function itself may, moreover, be random with an unknown distribution depending, of course, on θ and d.

(2) After each decision d_n is made, the decision-maker is told the value of the random loss incurred by him.

We are going to define a decision rule based on these assumptions and show that the resulting regret goes uniformly to zero. We will do

this for the two-decision case $D = \{1,2\}$ and the random loss L uniformly bounded; the result, however, readily extends to the case of D finite and L with uniformly bounded third moment (see [9] for the case of a game situation).

The decision rule suggested here is simple and more or less indicated by intuition. First, the sequence of indices n is divided into blocks J_k of increasing length and a net of countable partitions $\mathcal{B}_k$ of the sample space is defined. Before each decision d_n, a coin is flipped with probability of a head p_k if $n \in J_k (p_k \to 0$ as $k \to +\infty)$. The outcome of the toss determines whether the n-th step will be a test step (to gain information) or an active step (to minimize the loss). At test steps (when the outcome is a head), either decision is taken by random with equal probability. After the loss L_n is learned, the quantity $\psi_n = \pm(1 + L_n)$ is computed with plus sign if $d_n = 1$ and minus sign if $d_n = 2$. ψ_n as well as the sample X_n are remembered. At active steps within the block J_k, the decision $d_n = 1$ or 2 is taken according to whether the sum of the ψ_i's over all the past test steps in the same block J_k, at which the sample X_i fell into the same set of the partition $\mathcal{B}_k$ as the sample X_n just observed, is positive or negative. Loosely speaking, the decision is taken exhibiting smaller accumulated loss in the past test steps, where "nearly the same" sample was observed.

II. THE GENERIC DECISION PROBLEM

Throughout this paper $(\Omega,\mathcal{A},P)$ will always denote the basic probability space. The expectation of a random variable X will be denoted by $\mathbf{E}(X)$, the conditional expectation given a sub-σ-field $\mathcal{G}$ by $\mathbf{E}(X|\mathcal{G})$. The symbol $I_A(\cdot)$ will denote the indicator function of a set A. Other symbols or definitions will either be introduced later or used according to [4].

The basic decision problem generating the sequence is defined as follows.

Let $(\Xi,\mathfrak{X})$, the sample space, be any measurable space with the σ-field of subsets $\mathfrak{X}$ generated by a net of countable partitions $\{\mathcal{B}_k : k=1,2,\ldots\}$ of the set Ξ; that is, the partitions $\mathcal{B}_k = \{B_{jk} : j=1,2,\ldots\}$, where $\emptyset \neq B_{jk} \subset \Xi$; $j=1,2,\ldots$; $B_{j_1k} \cap B_{j_2k} = \emptyset$ for $j_1 \neq j_2$ and $\bigcup_{j=1}^{\infty} B_{jk} = \Xi$, satisfy the conditions:

(a) for every $A \in \mathcal{B}_k$ there exists $B \in \mathcal{B}_{k-1}$ such that $A \subset B$;
(b) $\mathfrak{X}$ is the minimum σ-field over the sequence $\{\mathcal{B}_k : k=1,2,\ldots\}$.

(2.1)

Notice that all the sets B_{jk} are therefore $\mathfrak{X}$-measurable.

Let Θ be an arbitrary set, the parameter space, and let $\mathcal{P} = \{\nu_\theta : \theta \in \Theta\}$ be a family of probability measures on the space $(\Xi,\mathfrak{X})$. We will assume that $\mathcal{P}$ is dominated by a σ-finite measure μ on $(\Xi,\mathfrak{X})$, i.e., that

$$\nu_\theta \ll \mu \qquad \text{for every} \quad \theta \in \Theta , \tag{2.2}$$

so that the Radon-Nikodym derivatives

$$\mathfrak{f}_\theta(x) = \frac{d\nu_\theta(x)}{d\mu(x)} , \; x \in \Xi ,$$

exist for every $\theta \in \Theta$.

Next, let $D = \{1,2\}$ be the decision space and let L be a random loss function defined on $\Theta \times D$, i.e., for every $\theta \in \Theta$, $d \in D$, $L(\theta,d)$ is a real-valued random variable. We will assume that L is uniformly bounded--more precisely, that there is a finite constant further on denoted by C such that

$$|L(\theta,d) + 1| \leq C \qquad \text{for every} \quad \theta \in \Theta,\ d \in D\ . \tag{2.3}$$

Let

$$0 \leq \ell(\theta,d) = \mathbf{E}\{L(\theta,d)\}\ , \qquad \theta \in \Theta,\ d \in D\ . \tag{2.4}$$

To avoid the trivial case, we will also assume that

$$\ell(\theta,1) \neq \ell(\theta,2) \qquad \text{for at least one} \quad \theta \in \Theta\ . \tag{2.5}$$

Let $\mathcal{J}$ be a σ-field of all subsets of Θ and let T be the class of all purely atomic finite signed measures on $(\Theta,\mathcal{J})$ with a finite number of atoms. Let $T_o \subset T$ be the subclass of all probability measures. If $\tau \in T_o$ has a single atom $\{\theta\}$, we will sometimes write simply θ instead of τ. Further, we will denote

$$\ell(\tau,d) = \int_\Theta \ell(\theta,d)\ d\tau(\theta)\ , \qquad \tau \in T,\ d \in D\ ,$$

and

$$\Delta(\tau) = \ell(\tau,2) - \ell(\tau,1)\ .$$

Notice that since $\tau \in T$ the above integral exists and by (2.3) is uniformly bounded by C. Similarly, let

$$w(x,\tau,d) = \int_\Theta \ell(\theta,d)\ \mathcal{F}_\theta(x)\ d\tau(\theta)\ , \qquad x \in \Xi,\ \tau \in T,\ d \in D\ ,$$

where, clearly $w(\cdot,\tau,d)$ is $\mathfrak{X}$-measurable and μ-integrable for every $\tau \in T$ and $d \in D$. Let

$$\rho(\tau) = \int_{\Xi} \min_{d \in D} \{w(x,\tau,d)\} \, d\mu(x)$$

be the Bayes risk for the hypothetical prior $\tau \in T$. Notice that, with addition and multiplication by a constant defined naturally on T, both $\ell(\tau,d)$ and $\Delta(\tau)$ are linear in τ and that ρ is a nonnegative concave continuous functional on T bounded by C on T_o.

III. THE SEQUENTIAL COMPOUND DECISION PROBLEM

The sequential compound decision problem is obtained by infinite repetition of the generic decision problem with $\theta \in \Theta$ varying arbitrarily along the sequence.

Let Θ^∞ be the set of all sequences

$$\boldsymbol{\theta} = \{\theta_n : n=0,1,\ldots\} .$$

For convenience, we will assume that θ_o is such that $L(\theta_o,\cdot) = 0$. There is no loss of generality in this assumption since the sequence in fact begins with θ_1, and θ_o is merely a dummy parameter.

Given a sequence $\boldsymbol{\theta}$ and an interval of integers $J = (n_1,n_2]$, $0 \leq n_1 < n_2$, we define $\tau_J \in T_o$ by

$$\tau_J(A) = (n_2-n_1)^{-1} \sum_{n=n_1+1}^{n_2} I_A(\theta_n) , \qquad A \in \tau . \tag{3.1}$$

If $J = (0,n]$, we will write simply τ_n instead of $\tau_{(0,n]}$.

Next, let $\{X_n : n=0,1,\ldots\}$ be a sequence of independent random variables taking values in $(\Xi,\mathfrak{X})$ with

$$PX_n^{-1} = \nu_{\theta_n} , \qquad \theta_n \in \boldsymbol{\theta} .$$

and let $\{L_n(\theta_n,d_n) : \theta_n \in \theta,\ d_n \in D\}$ be a sequence of independent random variables distributed as $L(\theta,d)$ whenever $\theta_n = \theta$, $d_n = d$. The sequences $\{X_n : n=0,1,\ldots\}$ and $\{L_n(\theta_n,d_n) : n=0,1,\ldots\}$ are assumed to be mutually independent. Here, X_n is the observed sample and $L_n(\theta_n,d_n)$ is the random loss for the decision d_n at n-th step.

Let s^* be a mapping from $(-\infty, +\infty)$ into D defined by:

$$s^*(y) = \begin{cases} 1 & \text{if } y > 0 , \\ \text{arbitrary} & \text{if } y = 0 , \\ 2 & \text{if } y < 0 . \end{cases}$$

We will need the following

Lemma. Let $\{U_n' : n=0,1,\ldots\}$ and $\{V_n : n=0,1,\ldots\}$ be mutually independent sequences of Bernoulli random variables with

$$P\{U_n' = 1\} = p , \qquad 0 < p \leq \frac{1}{2} ,$$

and

$$P\{V_n = 1\} = P\{V_n = 0\} = \frac{1}{2} ;$$

let

$$S_n = s^*\left(\sum_{i=0}^{n} Y_i'\right) ,$$

where

$$Y_n' = 2p^{-1}U_n'[V_n(1+L_n(\theta_n,2)) - (1-V_n)(1+L_n(\theta_n,1))] .$$

Then there is a finite constant C_1 depending only on the constant C in (2.3) such that for every $n=1,2,\ldots$, $\boldsymbol{\theta} \in \Theta^\infty$,

$$\mathbf{E}\left\{n^{-1}\sum_{i=1}^{n} l(\theta_i,S_i)\right\} - \min_{d \in D}\{l(\tau_n,d)\} \leq C_1(pn)^{-1/2} .$$

Proof.

Let $\tilde{Y}_n = Y_n' - \mathbf{E}\{Y_n'\}$. Notice that $\mathbf{E}\{Y_n'\} = \Delta(\theta_n)$ and that the random variables $\tilde{Y}_n$ are uniformly bounded,

$$|\tilde{Y}_n| \le cp^{-1} , \tag{3.2}$$

where $c < +\infty$ is a constant independent of p and θ_n. Further, elementary computation yields

$$\mathbf{E}|\tilde{Y}_n|^r \le (4C)^r p^{1-r} , \qquad r = 1,2,3 , \tag{3.3}$$

and for $0 < p \le ½$

$$4p^{-1} \le \mathbf{E}|\tilde{Y}_n|^2 . \tag{3.4}$$

Let

$$Z_n = n^{-½} \sum_{i=0}^{n} \tilde{Y}_i , \qquad n = 1,2,\ldots ;$$

let F_n denote the distribution function of the law $\mathfrak{L}(Z_n)$. Since the random variables $\tilde{Y}_n$ are independent, are centered at expectations, and have positive variance, the Berry-Esseen normal approximation (see [4], p. 288) yields

$$\sup_{x \in (-\infty,+\infty)} |F_n^*(x) - G^*(x)| \le \beta \left(\sum_{i=0}^{n} \mathbf{E}|\tilde{Y}_i|^2\right)^{-3/2} \sum_{i=0}^{n} \mathbf{E}|\tilde{Y}_i|^3 , \tag{3.5}$$

where β is the Berry-Esseen constant, G^* is the distribution function of the normal law $\mathfrak{N}(0,1)$, and F_n^* is the distribution function of the law

$$\mathfrak{L}\left[\left(\sum_{i=0}^{n} \mathbf{E}|\tilde{Y}_i|^2\right)^{-1/2} \sum_{i=0}^{n} \tilde{Y}_i\right] .$$

Denote

$$\bar{\sigma}_n = \left(n^{-1} \sum_{i=0}^{n} \mathbb{E}|\tilde{Y}_i|^2\right)^{-1/2} ; \qquad n = 1,2, \ldots .$$

Since,

$$\left(\sum_{i=0}^{n} \mathbb{E}|\tilde{Y}_i|^2\right)^{-3/2} \sum_{i=0}^{n} \tilde{Y}_i = \bar{\sigma}_n^{-1} Z_n$$

we have

$$F_n^*(x) = F_n(\bar{\sigma}_n x) .$$

Further by (3.3) and (3.4),

$$\left(\sum_{i=0}^{n} \mathbb{E}|\tilde{Y}_i|^2\right)^{-3/2} \sum_{i=0}^{n} \mathbb{E}|\tilde{Y}_i|^3 \leq 8C^3(n+1)^{-1/2}p^{-1/2} ,$$

so that (3.5) becomes

$$\sup_{x \in (-\infty,+\infty)} \left|F_n(x) - G^*\left(\bar{\sigma}_n^{-1} x\right)\right| \leq 8C^3\beta(n+1)^{-1/2}p^{-1/2}; \qquad n = 1,2, \ldots .$$

Hence for any real numbers $x_1, x_2; x_1 \leq x_2$;

$$P\{x_1 \leq Z_n < x_2\} \leq (x_2-x_1) \left(\frac{n}{n+1}\right)^{1/2} p^{1/2}$$

$$+ 16C^3\beta(n+1)^{-1/2}p^{-1/2} ; \qquad n = 1,2, \ldots , \tag{3.6}$$

since

$$G^*\left(\bar{\sigma}_n^{-1} x_2\right) - G^*\left(\bar{\sigma}_n^{-1} x_1\right) \leq (x_2-x_1)\bar{\sigma}_n^{-1} \leq (x_2-x_1) \left(\frac{n}{n+1}\right)^{1/2} p^{1/2}$$

by (3.4).

So equipped, we can start proving the lemma. Let us denote the left-hand side of the inequality to be proved by

$$Q_n = \mathbf{E}\left\{n^{-1} \sum_{i=1}^{n} \ell(\theta_i, s_i)\right\} - \varphi(\tau_n) ,$$

where

$$\varphi(\tau_n) = \min_{d \in D} \{\ell(\tau_n, d)\} .$$

We have

$$Q_n = \mathbf{E}\left\{n^{-1} \sum_{i=1}^{n} [i\ell(\tau_i, s_i) - (i-1)\ell(\tau_{i-1}, s_i)]\right\} - \varphi(\tau_n)$$

$$= \mathbf{E}\left\{n^{-1}\left[\sum_{i=1}^{n} i\ell(\tau_i, s_i) - \sum_{i=1}^{n-1} i\ell(\tau_i, s_i) - n\varphi(\tau_n)\right]\right\}$$

$$= \mathbf{E}\left\{n^{-1} \sum_{i=1}^{n-1} i[\ell(\tau_i, s_i) - \ell(\tau_i, s_{i+1})] + n[\ell(\tau_n, s_n) - \varphi(\tau_n)]\right\} , \tag{3.7}$$

where $\tau_i \in T_o$ is defined by (3.1).

Before going further, notice that for $\tau \in T$, x a real number, $d_1 = s^*(\Delta(\tau) + x)$ and $d_2 \in D$ we have the inequality

$$\ell(\tau, d_1) - \ell(\tau, d_2) \leq x(d_2 - d_1) . \tag{3.8}$$

To see this, let $t_x \in T$ be such that $\Delta(t_x) = x$. For example, t_x may have a single atom $\{\theta\}$ with $t_x(\{\theta\}) = x[\Delta(\theta)]^{-1}$ for some $\theta \in \Theta$ such that $\Delta(\theta) \neq 0$ (see assumption 2.4). Now, since $\Delta(\tau+t_x) = \Delta(\tau) + x$, we have $\varphi(\tau+t_x) = \ell(\tau+t_x,d_1) \leq \ell(\tau+t_x,d_2)$ so that by linearity of $\ell(\cdot,d)$

$$\ell(\tau,d_1) - \ell(\tau,d_2) \leq \ell(t_x,d_2) - \ell(t_x,d_1)$$

$$= \begin{cases} \Delta(t_x) & \text{if } d_1 = 1,\ d_2 = 2\ ; \\ -\Delta(t_x) & \text{if } d_1 = 2,\ d_2 = 1; \\ 0 & \text{if } d_1 = d_2\ ; \end{cases}$$

which together with $\Delta(t_x) = x$ proves (3.8).

We apply this inequality to the summands in (3.7). Since by definition

$$S_n = s^*\left(\sum_{i=1}^{n} Y_i'\right) = s^*\left(n^{-1} \sum_{i=0}^{n} Y_i'\right)$$

and

$$n^{-1} \sum_{i=0}^{n} Y_i' = n^{-1} \sum_{i=0}^{n} \mathbb{E}\{Y_i'\} + n^{-1} \sum_{i=0}^{n} \tilde{Y}_i = \Delta(\tau_n) + n^{-\frac{1}{2}} Z_n\ ; \qquad n = 1,2,\ldots\ ;$$

we can set $x = i^{-\frac{1}{2}} Z_i$ and $d_2 = S_{i+1}$ in (3.8), thus obtaining

$$\ell(\tau_i,S_i) - \ell(\tau_i,S_{i+1}) \leq i^{-\frac{1}{2}} Z_i\,(S_{i+1}-S_i)\ . \tag{3.9}$$

Similarly, setting $d_2 = s^*[\Delta(\tau_n)]$ and using the fact that by the definition of the mapping s^*, $\ell(\tau,d_2) = \varphi(\tau)$, we obtain

$$\ell(\tau_n, s_n) - \varphi(\tau_n) \leq n^{-1/2} Z_n [s^*(\Delta(\tau_n)) - s_n] \tag{3.10}$$

Substituting (3.9) and (3.10) into (3.7), we have

$$Q_n \leq \mathbf{E}\left\{ n^{-1} \sum_{i=1}^{n-1} i^{1/2} Z_i \, (s_{i+1} - s_i) + n^{1/2} Z_n [s^*(\Delta(\tau_n)) - s_n] \right\}$$

$$= n^{-1} \sum_{i=1}^{n} \mathbf{E}\left\{ [(i-1)^{1/2} Z_{i-1} - i^{1/2} Z_i] s_i \right\} + n^{-1/2} \, \mathbf{E}\{Z_n\} s^*(\Delta(\tau_n)) \ ,$$

where $Z_o = 0$. However, by definition, $(i-1)^{1/2} Z_{i-1} - i^{1/2} Z_i = -\tilde{Y}_i$ and since Y_i's are centered at expectations $\mathbf{E}\{Z_n\} = 0$. Therefore,

$$Q_n \leq -n^{-1} \sum_{i=1}^{n} \mathbf{E}\{\tilde{Y}_i s_i\} \ . \tag{3.11}$$

Next, by definition of s_i and the fact that $\mathbf{E}\{\tilde{Y}_i\} = 0$,

$$-\mathbf{E}\{\tilde{Y}_i s_i\} = - \int_{\left\{\sum_{j=0}^{i} Y'_j \geq 0\right\}} \tilde{Y}_i dP - 2 \int_{\left\{\sum_{j=0}^{i} Y'_j < 0\right\}} \tilde{Y}_i dP = \int_{\left\{\sum_{j=0}^{i} Y'_j \geq 0\right\}} \tilde{Y}_i dP \ , \tag{3.12a}$$

or

$$-\mathbf{E}\{\tilde{Y}_i s_i\} = - \int_{\left\{\sum_{j=0}^{i} Y'_j > 0\right\}} \tilde{Y}_i dP - 2 \int_{\left\{\sum_{j=0} Y'_j \leq 0\right\}} \tilde{Y}_i dP = \int_{\left\{\sum_{j=0} Y'_j > 0\right\}} \tilde{Y}_i dP \ , \tag{3.12b}$$

depending on whether $S_i = 1$ or 2 when the argument of the mapping s^* is zero. Let us consider the former case first. We have

$$\left\{\sum_{j=0}^{i} Y'_j \geq 0\right\} = \left\{\sum_{j=0}^{i-1} \tilde{Y}_j \geq -\tilde{Y}_i - \sum_{j=0}^{i} \Delta(\theta_j)\right\},$$

and by (3.2)

$$H'_i \subset \left\{\sum_{j=0}^{i-1} \tilde{Y}_j \geq -\tilde{Y}_i - \sum_{j=0}^{i} \Delta(\theta_j)\right\} \subset H''_i ,$$

where

$$H'_i = \left\{\sum_{j=0}^{i-1} \tilde{Y}_j \geq cp^{-1} - \sum_{j=0}^{i} \Delta(\theta_j)\right\},$$

and

$$H''_i = \left\{\sum_{j=0}^{i-1} \tilde{Y}_j \geq -cp^{-1} - \sum_{j=0}^{i} \Delta(\theta_j)\right\}.$$

Hence

$$\int_{\left\{\sum_{j=0}^{i} Y'_j \geq 0\right\}} \tilde{Y}_i dp = \int_{\left\{\sum_{j=0}^{i} Y'_j \geq 0\right\}} \tilde{Y}_i dP - \int_{\left\{\sum_{j=0}^{i} Y'_j \geq 0\right\}} \tilde{Y}_i^- dP \leq \int_{H''_i} Y_i^+ dP - \int_{H'_i} \tilde{Y}_i^- dP$$

$$= \int_{H_i} \tilde{Y}_i^+ dP + \int_{H'_i} \tilde{Y}_i dP ,$$

where $H_i = H''_i - H'_i$.

However, since the random variables $\tilde{Y}_i$ are independent,

$$\int_{H_i} Y_i^+ dP = P(H_i) \int_\Omega Y_i^+ dP \le P(H_i) \mathbf{E}|\tilde{Y}_i| \le 4CP(H_i)$$

by (3.3) and

$$\int_{H_i'} \tilde{Y}_i dP = P(H_i') \mathbf{E}\{\tilde{Y}_i\} = 0 \ .$$

$$H_i = \left\{ -c(i-1)^{-\frac{1}{2}} p^{-1} - (i-1)^{-\frac{1}{2}} \sum_{j=0}^{i} \Delta(\theta_j) \le Z_{i-1} < c(i-1)^{-\frac{1}{2}} p^{-1} - (i-1)^{-\frac{1}{2}} \sum_{j=0}^{i} \Delta(\theta_j) \right\} ,$$

so that by (3.6)

$$P(H_i) \le 2c(ip)^{-\frac{1}{2}} + 16C^3 \beta (ip)^{-\frac{1}{2}}$$

for $i = 2,3,\ldots$ and trivially also for $i = 1$. Thus (3.12a) becomes

$$- \mathbf{E}\{\tilde{Y}_i S_i\} \le (8Cc + 64C^4 \beta)(ip)^{-\frac{1}{2}} \ . \tag{3.13}$$

For the case (3.12b), it is easy to see that exactly the same reasoning applies so that (3.13) holds again. Substituting (3.13) into (3.11), we have finally

$$Q_n \le (8Cc + 64C^4 \beta) p^{-\frac{1}{2}} n^{-1} \sum_{i=1}^{n} i^{-\frac{1}{2}} ,$$

which together with $n^{-1} \sum_{i=1}^{n} i^{-\frac{1}{2}} \le \frac{3}{2} n^{-\frac{1}{2}}$ proves the lemma.

IV. THE DECISION RULE

Let $\{M_k : k=1,2,\ldots\}$ be a sequence of positive integers; let $\{N_k : K=0,1,\ldots\}$ be a sequence of integers defined by

$$N_k = N_{k-1} + M_k, \qquad k = 1,2,\ldots , \qquad N_o = 0;$$

let $J_k = (N_{k-1}, N_k]$ be intervals of integers.

Let $\{U_n : n=0,1,\ldots\}$ be a sequence of independent random variables taking values 0 and 1 with $P\{U_n = 1\} = p_k$ whenever $n+1 \in J_k$, let $\{V_n : n=0,1,\ldots\}$ be a sequence of i.i.d. random variables, taking values 0 and 1 with probability ½.

The sequences $\{U_n\}$ and $\{V_n\}$ are supposed to be mutually independent and also independent on the sequences $\{X_n\}$ and $\{L_n\}$ defined in Section III. The random variable U_n determines whether the n-th step will be a test step $(U_n = 1)$ or an active step $(U_n = 0)$, and V_n determines whether the decision $d = 1$ $(V_n = 1)$ will be used if the n-th step is a test one. Let $\boldsymbol{\theta} = \{\theta_n : n=0,1,\ldots\} \in \Theta^\infty$ be a sequence of parameters and let

$$\psi_n = V_n[1 + L_n(\theta_n,2)] - (1-V_n)[1 + L_n(\theta_n,1)] \ .$$

The sequences $\{\psi_n\}$, together with the sequence of samples $\{X_n\}$, represent the entire information the decision-maker is receiving.

The decision rule is now defined as follows:

At the n-th step, $n \in J_k$; $k=1,2, \ldots$;

$$\left.\begin{array}{l} 1/ \text{ if } U_n = 1 \text{ then decide } d_n^* = V_n + 1; \\ 2/ \text{ if } U_n = 0 \text{ and } X_n \in B_{jk} \text{ then decide} \\ \qquad d_n^* = s^*\left[\sum_{i=N_{k-1}}^{n-1} U_i \psi_i I_{B_{jk}}(X_i)\right] . \end{array}\right\} (*)$$

Theorem. Let $\{d_n^* : n=1,2,\ldots\}$ be the sequence of decisions resulting from the decision rule (*), where $p_k \to 0$ and $M_k p_k \to +\infty$ as $k \to +\infty$, let the assumptions (2.1) - (2.5) be satisfied. Then

$$\limsup_{n\to+\infty} \mathbf{E}\left\{N^{-1} \sum_{n=1}^{N} L_n(\theta_n, d_n^*) - \rho(\tau_N)\right\} = 0$$

uniformly in all sequences $\boldsymbol{\theta} \in \Theta^\infty$.

Proof.

Let $\boldsymbol{\theta}$ be a sequence from Θ^∞. Since, by (*), d_n^* and the random loss L_n are independent $\mathbf{E}\{L_n(\theta, d_n^*)\} = \mathbf{E}\{\ell(\theta_n, d_n^*)\}$. Let

$$Y_n = 2p_k^{-1} U_n \psi_n, \quad n+1 \in J_k, \quad k = 1,2, \ldots, \tag{4.1}$$

and let

$$S_n = s^*\left(\sum_{i=N_{k-1}}^{n} Y_i I_{B_{jk}}(X_i)\right), \tag{4.2}$$

whenever

$$X_n \in B_{jk}, \quad n \in J_k, \quad k = 1,2, \ldots .$$

We have

$$\mathbf{E}\left\{N^{-1} \sum_{n=1}^{N} L_n(\theta_n, d_n^*) - \rho(\tau_n)\right\} = \mathbf{E}\left\{N^{-1} \sum_{n=1}^{N} [\ell(\theta_n, d_n^*) - \ell(\theta_n, S_n)]\right\}$$

$$+ \mathbf{E}\left\{N^{-1} \sum_{n=1}^{N} \ell(\theta_n, S_n) - \rho(\tau_N)\right\}. \tag{4.3}$$

Now, by $(*)$, (4.1), and (4.2), $U_n = 0$, $X_n \in B_{jk} \Longrightarrow d_n^* = S_n$, so that d_n^* can differ from S_n only if $U_n = 1$. Therefore

$$|\ell(\theta_n, d_n^*) - \ell(\theta_n, S_n)| \leq 2CU_n$$

whence

$$\mathbf{E}\left\{N^{-1} \sum_{n=1}^{N} [\ell(\theta_n, d_n^*) - \ell(\theta_n, S_n)]\right\} \leq 2C\mathbf{E}\left\{N^{-1} \sum_{n=1}^{N} U_n\right\} . \tag{4.4}$$

However $\mathbf{E}\{U_n\} = p_k$ for $(n+1) \in J_k$ so that the bound in (4.4) and therefore also the first term on the right-hand side of (4.3) go to zero as $k \to +\infty$. For the second term, let us denote

$$R_{(n_1,n_2]} = \mathbf{E}\left\{(n_2-n_1)^{-1} \sum_{n=n_1+1}^{n_2} \ell(\theta_n, S_n)\right\} - \rho(\tau_{(n_1,n_2]}) ,$$

where $0 \leq n_1 < n_2$ are integers and let $N \in J_K$ for some $K = 1,2, \ldots$. Since ρ is concave on T_o, we have the inequality

$$\rho(\tau_N) \geq \sum_{k=1}^{k-1} N^{-1} M_k \rho(\tau_{J_k}) + N^{-1}(N-N_{k-1}) \, \rho(\tau_{(N_{k-1},N]}) ,$$

which implies

$$NR_N \leq \sum_{k=1}^{k-1} M_k R_{J_k} + (N-N_{k-1}) \, R_{(N_{k-1},N]} ,$$

where we wrote for short R_N instead of $R_{(0,N]}$.

We are going to prove that

$$\limsup_{k \to +\infty} R_{J_k} = 0 \qquad \text{uniformly in } \theta \in \Theta^\infty , \tag{4.5}$$

which in turn implies

$$\limsup_{n \to +\infty} R_N = 0 \qquad \text{uniformly in } \theta \in \Theta^{\infty}, \tag{4.6}$$

To see this, let $\epsilon > 0$. Since $M_k \to +\infty$ as $k \to +\infty$, there exists some $k(\epsilon)$ such that

$$M_k \geq (N-N_{k-1}) > M_{k(\epsilon)} \Longrightarrow R_{(N_{k-1},N]} < \epsilon$$

whatever be the sequence $\theta \in \Theta^{\infty}$. Hence, since

$$R_{(N_{k-1}),N]} \leq 2C,$$

we have for K large enough

$$N^{-1}\left[\sum_{k=1}^{k(\epsilon)} M_k R_{J_k} + \sum_{k=k(\epsilon)+1}^{k-1} M_k R_{J_k} + (N-N_{k-1})\, R_{(N_{k-1},N]}\right]$$

$$< N^{-1}\left[2C \sum_{k=1}^{k(\epsilon)} M_k + \sum_{k=k(\epsilon)+1}^{k-1} M_k + (N-N_{k-1})q\right], \tag{4.7}$$

where $q \leq 2C$ if $N - N_{k-1} \leq M_{k(\epsilon)}$ and $q < \epsilon$ if $N - N_{k-1} > M_{k(\epsilon)}$. However

$$N = \sum_{k=1}^{k-1} M_k + (N-N_{k-1})$$

so that the first term on the right-hand side of (4.7) as well as the last one in the case $N - N_{k-1} \leq M_{k(\epsilon)}$ can be made arbitrarily small while the second term and the last one in the case $N - N_{k-1} > M_{k(\epsilon)}$ are both smaller than ϵ. This proves (4.6).

Let us now start proving (4.5). To simplify notation, let us relabel the subscripts at θ_n and S_n, writing

$$R_{J_k} = E\left\{M_k^{-1} \sum_{n=1}^{M_k} \hat{\ell}(\theta_n, S_n)\right\} - \rho(\tau_{J_k}) .$$

Let

$$\rho_k(\tau) = \sum_{j=1}^{\infty} \min_{d \in D} \int_{B_{jk}} \int_{\Theta} (\theta, d) f_\theta(x) d\tau(\theta) d\mu(x) \ , \ \tau \in T_o, \ k = 1,2,\ldots \ , \tag{4.8}$$

and let

$$w_k(x,\tau,d) = \int_\Theta \ell(\theta,d) f_\theta^{(k)}(x) d\tau(\theta) .$$

where

$$f_\theta^{(k)}(x) = \sum_j \nu_\theta(B_{jk})[\mu(B_{jk})]^{-1} I_{B_{jk}}(x); \ x \in \Xi ,$$

the summation being over those j's for which $\mu(B_{jk}) > 0$. Clearly, $w_k(\cdot,\tau,d)$ is $\mathfrak{X}$-measurable and μ-integrable for every $\tau \in T_o$, $d \in D$ and

$$\rho_k(\tau) = \int_\Xi \min_{d \in D} \{w_k(x,\tau,d)\} \, d\mu(x) .$$

Let $\mathcal{F}_k$ be the minimum σ-field over the partition $\mathcal{B}_k$. Since by the assumption (2.1) $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \subset \ldots \subset \mathfrak{X}$. $\{f_\theta^{(k)} : k=1,2, \ldots\}$ is for every $\theta \in \Theta$ a martingale sequence on the measure space $(\Xi,\mathfrak{X},\mu)$ closed on the right by f_θ. Moreover, f_θ is the nearest $\mathfrak{X}$-measurable μ-integrable function closing $\{f_\theta^{(k)} : k=1,2,\ldots\}$ on the right. This can be

easily seen since if g is another such a function closing the sequence on the right, i.e., if

$$B \in \mathfrak{F}_k \Longrightarrow \int_B g d\mu = \int_B \mathfrak{F}_\theta^{(k)} d\mu$$

for all $k = 1, 2, \ldots,$ then also

$$B \in \mathfrak{A} \Longrightarrow \int_B g d\mu = \int_B \mathfrak{F}_\theta d\mu$$

since $\mathfrak{F}_k \uparrow +\infty$ as $k \to +\infty$ by (2.1). Hence the martingale closure theorem (see [4], p. 394) applies, yielding

$$\lim_{k \to +\infty} f_\theta^{(k)} \overset{\mu}{\equiv} f_\theta$$

for every $\theta \in \Theta$. Since $\tau \in T_o$ has a finite number of atoms, this implies that also

$$\lim_{k \to +\infty} \min_{d \in D} \{w_k(\cdot, \tau, d)\} \overset{\mu}{\equiv} w(\tau, d) ,$$

and since

$$\int_\Xi |w_k(x, \tau, d)| d\mu(x); \; k = 1, 2, \ldots ;$$

are uniformly bounded, it follows that for every $\tau \in T_o$

$$\lim_{k \to +\infty} \rho_k(\tau) = \rho(\tau) .$$

However $\rho, \rho_1, \rho_2, \ldots$ are continuous functionals on the compact T_o so that the above convergence is uniform in $\tau \in T_o$. Therefore

$$\lim_{k \to +\infty} |\rho_k(\tau_{J_k}) - \rho(\tau_{J_k})| = 0$$

uniformly in all sequences $\boldsymbol{\theta} \in \Theta$ and, in view of (4.5), it remains to prove that

$$\limsup_{k \to +\infty} R'_{J_k} = 0 \tag{4.9}$$

uniformly in $\boldsymbol{\theta} \in \Theta^{\infty}$, where

$$R'_{J_k} = \mathbf{E}\left\{M_k^{-1} \sum_{n=1}^{M_k} \ell(\theta_n, s_n)\right\} - \rho_k(\tau_{J_k}) .$$

Let

$$\Lambda_{jk} = \{n \in \{1, \ldots, M_k\} : X_n \in B_{jk}\} , \qquad \lambda_{jk} = \sum_{n=1}^{M_k} I_{B_{jk}}(X_n)$$

and let $\mathfrak{M}_{jk}$; $j = 1,2, \ldots$; $k = 1,2, \ldots$; be the sub-σ-field of $\mathcal{A}$ induced by the family

$$\left\{I_{B_{jk}}(X_1), \ldots, I_{B_{jk}}(X_{M_k})\right\} .$$

Now

$$\mathbf{E}\left\{M_k^{-1} \sum_{n=1}^{M_k} \ell(\theta_n, s_n)\right\} = \mathbf{E}\left\{\sum_{j=1}^{\infty} M_k^{-1} \sum_{n \in \Lambda_{jk}} \ell(\theta_n, s_n)\right\}$$

$$= \sum_{j=1}^{\infty} \mathbf{E}\left\{M_k^{-1} \sum_{n \in \Lambda_{jk}} \ell(\theta_n, s_n)\right\} ,$$

where the series in the last term absolutely converges since the summands are bounded by $C\,\mathbf{E}\{\lambda_{jk}M_k^{-1}\}$, which sum up to the constant C. Next by (4.8),

$$\rho_k(\tau_{J_k}) = \sum_{j=1}^{\infty} \min_{d \in D} \left\{ M_k^{-1} \sum_{n=1}^{M_k} \ell(\theta_n,d)\ \mathbf{E}\{I_{B_{jk}}(X_n)\} \right\}$$

$$= \sum_{j=1}^{\infty} \min_{d \in D} \mathbf{E} \left\{ M_k^{-1} \sum_{n \in \Lambda_{jk}}^{\infty} \ell(\theta_n,d) \right\} \geq \sum_{j=1}^{\infty} \mathbf{E}\{\varphi(\tau_{\Lambda_{jk}})\} ,$$

where

$$\varphi(\tau) = \min_{d \in D} \{\ell(\tau,d)\} .$$

Thus

$$R'_{J_k} \leq \sum_{j=1}^{\infty} \mathbf{E} \left\{ M_k^{-1} \sum_{n \in \Lambda_{jk}}^{\infty} \ell(\theta_n,s_n) - \varphi(\tau_{\Lambda_{jk}}) \right\}$$

$$= \sum_{j=1}^{\infty} \mathbf{E} \left\{ \mathbf{E}\left[M_k^{-1} \sum_{n \in \Lambda_{jk}}^{\infty} \ell(\theta_n,s_n) \middle| \mathfrak{M}_{jk} \right] - \varphi(\tau_{\Lambda_{jk}}) \right\} .$$

Applying now the Lemma of Section 2 to the terms under the first expectation sign, we obtain

$$R'_{J_k} \leq \sum_{j=1}^{\infty} \mathbf{E}\left\{ C_1 \left(p_k \lambda_{jk} M_k^{-1} \right)^{-1/2} \right\} = C_1 \sum_{j=1}^{\infty} (M_k p_k)^{-1/2} \left[\nu_{\tau_{J_k}}(B_{jk}) \right]^{1/2} , \qquad (4.10)$$

where

$$\nu_\tau(\cdot) = \int_\Theta \nu_\theta(\cdot) d\tau(\theta) .$$

The summands in (4.10), however, are also bounded by

$$2C\,\mathbf{E}\left\{\lambda_{jk} M_k^{-1}\right\} = 2C\nu_{\tau_{J_k}}(B_{jk}) ,$$

which sums up to the constant 2C since $\nu_{\tau_{J_k}}$ is a probability measure. Thus the dominated convergence theorem applies to (4.10) and together with the assumption $(M_k p_k) \to +\infty$ yield (4.9).

<u>The theorem is proved</u>.

V. FINAL REMARKS

Let us make a few remarks concerning the problem. First notice that the decision rule suggested here does not make use of all the information available to the decision-maker since the information obtained when active decisions are made is disregarded. This has been done mainly for the sake of the proof; it is, nevertheless, conceivable that if the disregarded information were used the convergence might be faster. The rate of convergence itself would be worth investigating. It was shown under a similar assumption in a game situation [9] that $R_n = O(n^{-1/3})$ (for a special choice of M_k). Here, the rate would probably also depend on the partitions $\mathcal{B}_k$ and possibly on the family $\mathcal{P}$. Reference [9] also indicates that the convergence of average losses instead of risks may also be proven. However, the question of modifying the decision rule so as to use the whole of the past information obtained and not to begin at each block from a scratch remains open for the time being.

REFERENCES

1. N. Alēns, and T. M. Cover. Compound Bayes learning without a teacher (Abstract). Proc. 1st Annual Princeton Conf. on Information Theory and Systems, 109, Princeton, New Yersey, 1967.

2. S. Jílovec, and B. Šubert. Repetitive play of a game against nature, Aplikace Matematiky No. 5., Vol. 12 (1967), Prague, Czechoslovakia.

3. J. Hannan. The dynamic statistical decision problem when the component problem involves a finite number, m, of distributions (Abstract). Ann. Math. Statist. 26, 37-51, (1956).

4. M. Loève. Probability theory (3rd ed.), Van Nostrand, Princeton, New Jersey, 1963.

5. J. Van Ryzin. Sequential compound decision problem with finite loss matrix. Ann. Math. Statist. 37, 954-975, (1966).

6. J. Van Ryzin. Repetitive play in finite statistical games with unknown distributions. Ann. Math. Statis. 37, 976-994, (1966).

7. E. Samuel. Asymptotic solutions of the sequential compound decision problem. Ann. Math. Statist. 34, 1079-1094, (1963).

8. B. Šubert. An asymptotically optimal decision procedure. Transaction of the 4th Prague Conf. on Information Theory..., 573-588, Prague, Czechoslovakia 1965.

9. B. O. Šubert. Repetitive play of an unknown game against nature. SEL Technical Report No. 6151-3, December 1967, Stanford University, California.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Stanford Electronics Laboratories Stanford University Stanford, California	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

LEARNING WITH A LACK OF PRIOR DATA

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report No. 6151-4

5. AUTHOR(S) (First name, middle initial, last name)

Bruno O. Shubert

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
December 1967	31	9

8a. CONTRACT OR GRANT NO. Nonr-225(83), NR 373 360 b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) SEL-67-107 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

Reproduction in whole or in part is permitted for any purpose of the United States Government.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Joint Services Electronics Programs U.S. Army, U.S. Navy, and U.S. Air Force

13. ABSTRACT

DD FORM 1 NOV 65 1473

UNCLASSIFIED
Security Classification